



**MALAYSIAN JOURNAL OF LEARNING
AND INSTRUCTION**

<https://e-journal.uum.edu.my/index.php/mjli>

How to cite this article:

Anwar Farhan Mohamad Marzaini, Shahazwan Mat Yusoff, Noorhayati Zakaria, Mohamad Helme Basal, Wan Nurul Elia Haslee Sharil & Kaarthiyany Supramaniam. (2024). A discovery on Malaysian ESL teachers' adherence to policy mandates in classroom-based assessment practices. *Malaysian Journal of Learning and Instruction*, 21(1), 249-284. <https://doi.org/10.32890/mjli2024.21.1.9>

**A DISCOVERY ON MALAYSIAN ESL TEACHERS'
ADHERENCE TO POLICY MANDATES IN CLASSROOM-
BASED ASSESSMENT PRACTICES**

**¹Anwar Farhan Mohamad Marzaini, ²Shahazwan Mat Yusoff,
³Noorhayati Zakaria, ⁴Mohamad Helme Basal, ⁵Wan Nurul Elia
Haslee Sharil & ⁶Kaarthiyany Supramaniam**

¹Academy of Language Studies, Universiti Teknologi MARA,
Cawangan Pulau Pinang, Malaysia

^{1,5&6}Faculty of Education,

Universiti Teknologi MARA, Selangor, Malaysia

²Department of Curriculum and Instructional Technology,
University of Malaya, Malaysia

³Department of Language and Literacy Education,
University of Malaya, Malaysia

⁴Faculty of Sports Science and Recreation,
Universiti Teknologi MARA, Malaysia

¹*Corresponding email: anwarmarzaini@uitm.edu.my*

Received: 14/7/2023 Revised: 28/12/2023 Accepted: 4/1/2024 Published: 18/1/2024

ABSTRACT

Purpose – Although there exists a substantial body of research discussing the efficient implementation of classroom-based

assessment (CBA), there is a dearth of knowledge regarding its application within the context of English as a Second Language (ESL) education in Malaysia. The current study aimed to investigate the ESL teachers' practices in implementing CBA, with a specific emphasis on the alignment of their assessment procedures with the policy requirements.

Methodology – Hence, this research utilizes a qualitative approach, employing data triangulation through semi-structured interviews, classroom observations, and document analysis. Four ESL teachers from two public secondary schools in the northeast district of the Malaysian city of Pulau Pinang were selected to participate in the study. The four ESL teachers taught the same grade (Form 2) at the secondary level. The data collected were analyzed through thematic analysis procedures.

Findings – This research discovered that teachers demonstrate commendable practices in planning the CBA. They exhibited discernment by formulating instructional objectives in line with the curriculum document. However, a misalignment was observed in teachers' practices during the implementation phase. Despite employing various assessment methods, teachers predominantly leaned towards formal assessments, with less emphasis on student-centred approaches. Challenges surfaced when explicit learning objectives were not effectively communicated, and the use of standardized rubrics, was not tailored to students' diverse capabilities. At the reporting phase, concerns were raised about an excessive focus on administrative data recording. This study brought to light that teachers relied on the accumulation of students' grades and marks to determine the mastery levels, thus corroborating with the exam-oriented practices.

Significance – This study holds significant value for the field of ESL education in Malaysia by shedding light on the actual implementation of CBA and its alignment with policy directives. By providing information on how policies are implemented at the local level, contributes to the improvement of assessment policies that can be tailored to classroom realities.

Keywords: Classroom-based assessment, English Second Language, teachers' practices, policy implementation, language assessment, curriculum and instruction.

INTRODUCTION

Malaysia's education was notoriously known for its exam-centric culture, as numerous high-stakes examinations were prioritized within the system. This archaic exam-oriented culture prompted a heated debate among numerous scholars, who argued that it has hindered and prevented teachers from maximizing their students' abilities and potential in the learning process (Chin et al., 2019; Azli & Akmar, 2020). Considering this issue, the government has made significant efforts to revamp the current education system in the country. The government began reforming the education system in 2021 by eliminating high-stakes exams such as *Ujian Pencapaian Sekolah Rendah* (UPSR) and *Pentaksiran Tingkatan 3* (PT3) which were foregrounded at the primary and secondary school level. Thus, classroom-based assessment (CBA) was introduced to replace these high-stakes exams as a means of improving the assessment system and shifting away from the exam-centric culture. (Rethnasamy et al., 2021). The CBA is viewed as a more comprehensive assessment procedure because it emphasizes the combination of formative and summative culture under the School-Based Assessment (SBA) umbrella.

Since educational reform in Malaysia was primarily disseminated from the top down, stakeholders at the micro level, particularly teachers, must take drastic action to adapt to this shifting assessment climate. Taneri (2016) added that this type of reform has significant ramifications in requiring teachers to comprehend the changes to take appropriate and effective action. The teachers' roles in executing CBA extended to every level, beginning with planning, implementing, evaluating, and reporting, the assessment data in accordance with the policymakers' directives (Mohd Isa et al., 2020). To date, the implementation of CBA in English language education has been aligned with the Common European Framework of References (CEFR). The integration of CEFR into the CBA system aims to produce a culture of progressive assessment in which teachers have the autonomy to periodically monitor and chart their students' language progression based on a standard international descriptor (Sidhu et al., 2018). In light of this aim and objectives desired by policymakers, teachers are required to evaluate their students' language development through language skills such as reading, writing, listening, and speaking by incorporating grammatical knowledge and literature into the

evaluation process (Hopfenbeck, 2018). Further, Pellegrino, DiBello, and Goldman (2016) asserted that the implementation of CBA in the context of the English language necessitates the teacher's ability and autonomy to design an instrument that encourages students to demonstrate their language knowledge by aligning it with CEFR principles.

Despite the widespread use of classroom assessment in the Malaysian educational system, it continues to highlight the challenges faced by teachers, particularly when integrating new approaches into the actual classroom setting (Pastore, 2023). Numerous studies have indicated that many teachers lack proficiency in executing proper assessment procedures, resulting in hasty implementations at the grassroots level (Arumugham, 2020; Suppian et al., 2020; Yeh, 2021; Mohamad Marzaini et al., 2023). Mat Yusoff et al. (2023) discovered teachers are still depending to summative and standardized testing which is comparable to the exam-oriented culture in the previous system. Lumadi (2013) supported that teachers are incomprehensive to the policy interpretation, which inadvertently affect their practices in planning, executing, and grading the assessment. Their assessment methods were not diversified due to time constraints which has complicated the adherence to the curriculum needs (Lumadi, 2013). Panadero et al. (2019) study also unravelled that formative assessments like peer or self-assessments are not in favoured by most teachers as students' mastery levels are still determined by their final score in the exam. Worse, some quarters argue that the swift implementation of CBA after the abolition of PT3 in secondary schools by the government was viewed as a drastic move, contributing to these issues (Marnizam & Ali, 2021). In addition, the limited training and dissemination process for this new policy has exacerbated this issue, leaving teachers with significant resistance to adapting to the change (Ghavifekr et al., 2016). Thus, the expectation that teachers will master the new assessment policy remains hazy in light of this demand. Although the CBA is increasingly recognized as crucial to language learning (Liu & Xu, 2017), research on teachers' implementation of the CBA is relatively scant. The limited body of research has focused on certain aspects of CBA, such as using assessment methods and providing feedback (Yan et al., 2021; Chen et al., 2014; Cheng et al., 2008). There is a dearth of information in exploring the alignment of teachers' CBA practices with the policy need. Hence, this creates the urgency for a study such as this to fill the knowledge gap by discovering the alignment of teachers CBA practices with the curriculum need. By exploring

teachers' synchronisation to curriculum mandates, this study can better shed the light on the enactment of new assessment culture in Malaysian educational setting.

Hence, at the end, this study is intended to address the following research questions:

- 1) What strategies do ESL teachers in Malaysian secondary schools employ in implementing the classroom-based assessment?
- 2) To what degree do the assessment practices of Malaysian secondary school ESL teachers align with the guidelines outlined in the curriculum policy?

LITERATURE REVIEW

The Classroom-Based Assessment

Classroom-based assessment (CBA) aims to inform teacher decision-making and student learning (Black, 2015; Andrade & Cizek, 2010; Zhao et al., 2017). This assessment system demands teachers collaborative effort to gather data in determining if the learning has improved to make good decisions. Leong (2014) and Harlen (2012) highlighted that teachers are recommended to blend formative and summative assessments in the classroom, making the interaction a continuum. In this sense the formative evaluation adopts the theory of Vygotsky's Zone of Proximal Development. It helps teachers and students reach higher levels in classroom instruction (Black, 2015; Pattalitan Jr., 2016). Classroom assessment is increasingly curriculum-integrated and ongoing (Tee & Ahmed, 2014). Summative assessment, on the other hand, is "an accountability measure that is generally used as part of the grading process" (p. 1) and evaluates students' learning process at the end of each unit or chapter, term, or year. Many modern learning theories divide classroom evaluation into three categories: assessment of learning, assessment as learning, and assessment for learning (Pattalitan Jr., 2016). Malaysian classroom evaluations adhere to the same principles. "Assessment of learning" can be viewed as a form of summative assessment in which teachers evaluate students' performance on specific learning standards based on evidence. This paves the way for future educational programmes for students. "Assessment as learning" encourages students to reflect on their learning progress and to monitor their progress and learning

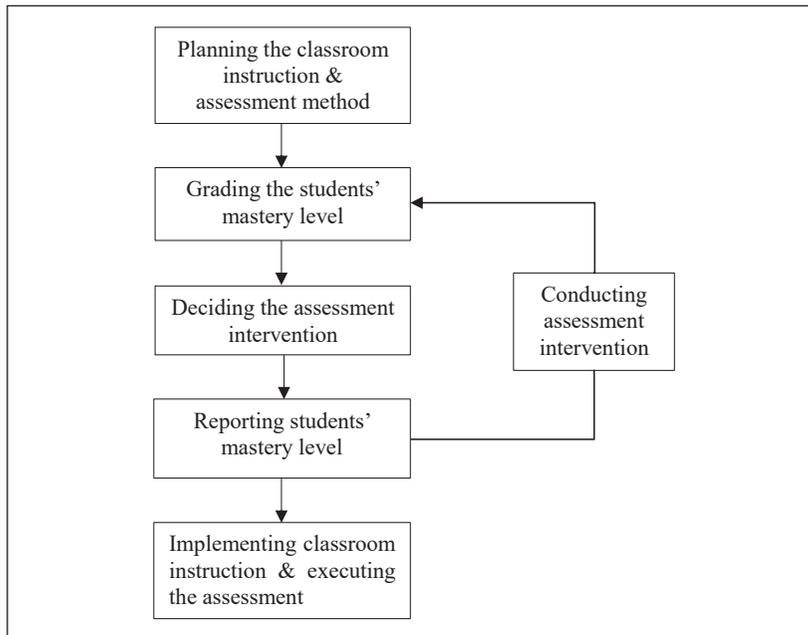
performance in order to attain higher standards. Students are entrusted with the responsibility of posing reflective queries and formulating multiple learning-enhancing strategies. Assessment for learning occurs when teachers provide students with constructive feedback based on inferences of student progress drawn from observations, anecdotes, question-and-answer sessions, and simple tests. This encourages and motivates students to enhance their learning.

The Implementation of Classroom-Based Assessment

The CBA is a iterative process which demands the teachers' full autonomy to navigate the flow of their classroom assessment. According to the Curriculum Development Division (2017), the CBA requires teachers to plan, construct items, administer, monitor, record, and report students' level of mastery in any subject taught when conducting classroom assessment. Teachers are needed to incorporate the CBA into their teaching and learning process in the classroom. Figure 1 illustrates the flow chart of CBA implementation.

Figure 1

Flow Chart of CBA Implementation



According to Figure 1, the implementation of CBA starts with the initial stage of developing the classroom instruction and evaluation technique. During this step, teachers are required to establish the specific learning objectives that they expect their students to achieve (Curriculum Development Division, 2017). They must also determine the assessment component that they will use to evaluate their students' progress. This assessment should be suited to the content and learning standards that have been identified. Therefore, this step also requires teachers to determine the evaluation procedures that are appropriate for the specific cultural setting of their classroom. The selection of this evaluation method relies on the learning objectives to be appraised and is suitable for the student's proficiency. Once the learning objectives have been determined, teachers can proceed with administering the assessment. Teachers can employ several approaches, including oral, written, and observational, to conduct continuous assessments in teaching and learning. During the implementation of CBA, teachers must ensure that the assessments are seamlessly incorporated into classroom instruction. Additionally, teachers should be granted the freedom to incorporate multiple assessment methods that align with the student's skills. During instructional activities, teachers assess a student's level of proficiency by evaluating their performance on topic and learning standard clusters. The teacher will assess the student's mastery level by evaluating the description of rubrics. This assessment will be recorded in a template and later analysed to determine the necessary steps for improving the student's developmental stage. The continuing evaluation, sometimes referred to as "assessment for learning," is crucial as it offers students prompt and pertinent feedback on how to enhance their work (Pattalitan Jr., 2016; Zhao et al., 2017; Tee & Ahmed, 2014). In addition, teachers must also execute the subsequent course of action, which might be either instantaneous or premeditated. The cognitive aptitude of each student to comprehend and excel in their subjects varies. During the assessment, teachers can identify pupils who have not yet or have fully acquired the knowledge they have been taught. Subsequently, the teacher might use suitable subsequent measures to enhance the degree of proficiency of each pupil. Finally, teachers must also compile a report regarding student's academic progress. Reporting involves periodically providing stakeholders, particularly parents, with information on students' progress, knowledge acquisition, skill mastery, value application, attitude development, and successes. By engaging in reporting, parents can effectively track their child's educational advancement and degree of expertise.

METHODOLOGY

Research Designs and Participant

In this study, a qualitative methodology was chosen as the primary approach to gather data pertaining to the implementation of CBA among ESL teachers. To do this, a case study design, as described by Yin (2014), was employed within a specific context. This design allowed for a comprehensive investigation of the phenomenon at hand by triangulating data from semi-structured interviews, document analysis, and classroom observation. Merriam and Tisdell (2016) advocated that the selection of a case study design was rooted in its suitability for in-depth exploration of the nuanced perspectives and intricacies associated with the subject of interest. It was considered a robust approach to simultaneously gain a thorough understanding of the phenomenon and provide a detailed description, aligning with Yin's (2014) framework. Additionally, Rahman (2014) argued convincingly in favour of incorporating the case study design as an invaluable tool for researchers. This methodology offers a unique opportunity to delve into the "lived experiences" of stakeholders involved in the process of implementing a new educational policy, such as CBA in this case. By studying the lived experiences of teachers in the new assessment climate, this study can better grasp the practical implications, challenges, and success associated with the adoption of such policies. In essence, the utilization of a case study design in this research served as a powerful means to comprehensively explore the real-world application of CBA by ESL teachers, uncovering not only the surface-level processes but also the deeper, contextual aspects of its implementation. This approach was well-founded in the literature, aligning with established principles of qualitative research and the need to comprehend the dynamic and multifaceted dimensions of educational policy implementation.

Therefore, a maximum variation sampling method (Merriam & Tisdell, 2016) was used to find participants with a wide range of demographics. For example, teachers who are different ages, teach at different grade levels, have different levels of education, and have been teaching for different amounts of time. This enables researchers to identify substantial shared patterns that transcend the sample's diversity. Consequently, four ESL teachers from two public secondary schools in the northeast district of the Malaysian city of Pulau Pinang were selected to participate in the study. The two schools differed in numerous respects. School A had approximately 70 classes, which is

double the number of classes at School B. School A was designated as the High-Performance School, where the processes of teaching, learning, and assessment were rigorously done. Regular research activities were also organized at School A to discuss how to enhance teaching by applying theories in the classroom. Since this research was qualitative, Creswell (2002) suggests that 4 to 25 participants would be sufficient to achieve the study's objectives. In the context of this study, these four ESL teachers taught the same grade (Form 2) at the secondary level. Moreover, the utilization of multiple data collection methods for each participant has indeed contributed to the in-depth exploration of the alignment of teachers' assessment practices in the implementation of CBA. As a result, this can shed more light on the implementation of CBA, as these participants have similar pedagogical and content knowledge pertaining to their field. In order to disperse the participants across a larger population, these teachers were chosen using the criterion-based selection method. Therefore, these teachers must have a bachelor's degree and at least three years of experience in ELT (English Language Teaching). The teachers must have attended the workshop and received training in relation to the CBA from the state or district education department and professional learning community (PLC) at the school level. The demographic profiles of the selected participants in this study are shown in Table 1. To maintain the privacy and confidentiality of the participants, a pseudonym was used for each participant.

Table 1

Participants' Demographic Profile

Participants	Gender	Age	Educational Background	Major	Teaching experience
Teacher 1	Female	35	B.Ed TESL	English	7
Teacher 2	Female	28	M.Ed TESL	Linguistic	4
Teacher 3	Female	30	B.Ed TESOL	English	6
Teacher 4	Female	27	B.Ed TESL	English	3

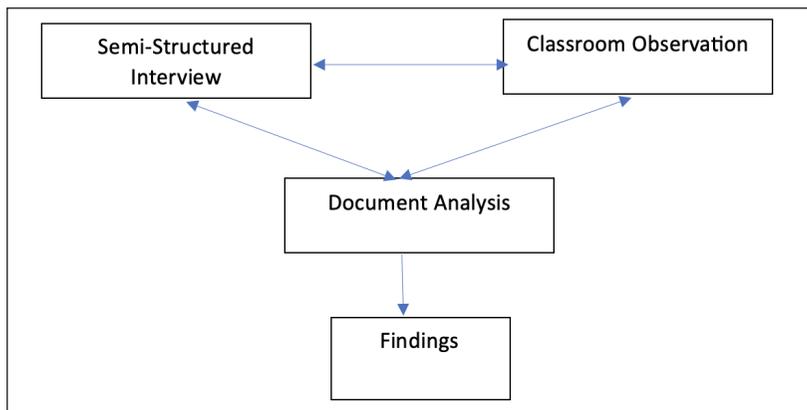
Data Collection

To ensure the reliability and validity of its findings, this study employed a multifaceted approach to data collection. As previously stated, the study utilized semi-structured interviews, classroom observations, and document analysis as essential data collection methods. Triangulation is a technique used to obtain a deeper

understanding of how teachers' assessment practices align with the objectives and requirements of the questioned educational policy. In this instance, classroom observations were used to further investigate the interview-derived data. This methodology enabled researchers not only to gain insights from teachers' verbal responses but also to observe their assessment practices in the actual classroom setting. It allowed for a broader and more contextual understanding of how these practices were implemented in practice. In addition, the interview and classroom observation data were meticulously cross-referenced and corroborated with the curriculum document analysis. This triangulation technique was instrumental in shedding light on the alignment of teachers' assessment practices with the educational policy's specific requirements and objectives. By cross-referencing these diverse sources, this method helps this study determine the extent to which teachers' actions and instructional decisions align with the policy's overarching goals (Rahman, 2014). The ultimate purpose of this triangulation strategy was to provide a comprehensive evaluation of teachers' compliance with policy requirements. This could then provide valuable insights into the prospective scenarios for improved policy implementation at the grassroots level, where the policy's practical implications are greatest. By utilizing multiple data sources and triangulation, the research aimed to increase the profundity and credibility of its findings, thereby contributing to a more robust analysis of teachers' assessment practices with the goals of educational policy. Figure 2 illustrates the triangulation procedures employed in this study.

Figure 2

Data Triangulation Procedures



The semi-structured interviews with open-ended questions served as the primary data collection technique for the implementation of CBA among ESL teachers. Creswell (2012) reaffirmed that the responses generated from open-ended inquiries can be a valuable instrument for researchers to collect insightful and precise information because it allows participants to respond in their own words and language. Further, Rahman (2014) claimed this method could encourage the participants' subjective ideas, thoughts, emotions, and perspectives and is a more effective way of shedding light on how the policy is implemented at the micro level. To ensure the consistency of the data collection process, an interview guide (Patton, 2015) was utilized to assist the researchers in keeping their attention on the predetermined subject of the study during the interview (Merriam, 1998). The interview instrument consists of three main sections. Section A of the interview deals with the teachers' personal background information which is related to their teaching experiences and educational background. Continuously, Section B of the interview deals with the teacher's knowledge of the implementation of CBA whereas, in Section C, teachers were asked to narrate their practices in executing the assessment in the classroom. The interview with the participants lasted between 20 to 30 minutes and it was tape-recorded for data analysis.

In addition, classroom observation was used to validate the teachers' interview data. The purpose of incorporating classroom observation into the current study was to ensure that specific information regarding teachers' CBA practices at the research site could be specified. This was supported by Merriam (1998), who asserted that the classroom observation technique could shed more light on the interaction between teacher and students in the classroom, which contributes to the implementation of the curriculum; therefore, an examination of classroom communication patterns is necessary to comprehend how the new policy is implemented in the classroom (p.127). Field notes were used to specify how teachers implemented the CBA to capture data from classroom observation. Two columns were created for the field notes. The first column was labelled 'Description of event' and contained a comprehensive explanation of the assessment practices of classroom instructors. The researchers meticulously described the classroom's physical environment, the interactions between students and teachers during assessment activities, and the assessment administration methods utilized during teaching and learning. The goal was to obtain a comprehensive and accurate depiction of

classroom assessment procedures so that the practices and methods of teachers could be analyzed in depth using dependable data (Rahman, 2014). The second column, 'Reflection,' provided space for the researcher to list the classroom procedures observed during the teacher observations. This made it easier to generalize the findings of the observation. Meticulous field notes were recorded throughout the observation, allowing for a complex and comprehensive description of the phenomenon under study (Merriam, 1998). The researchers also maintain the 'Complete Observer' position (Baker, 2006) when capturing data through classroom observation to minimize the 'Hawthorne effect' and ensure that participants engage in realistic assessment practises. Hence, there were no video recording of classroom instruction as it was intentionally omitted to allow teachers the freedom to authentically implement their assessment methods. The broad spectrum of teaching practices was observed among the teachers, with a particular emphasis on the development of diverse skills (listening, speaking, reading, writing). This observation aimed to explore the wide array of assessment techniques employed by teachers as they implemented the Classroom-Based Assessment (CBA) approach.

The teachers' assessment practices identified through interviews and classroom observation were then compared to the requirements of the curriculum. As a result, the Standard Based English Language Curriculum (SBELC) and the CBA Guidebook were collected and analysed via document analysis. Teachers must refer to the SBELC's curriculum standards and learning objectives when incorporating pedagogical practises into their CBA. Researchers utilised the document summary form (Yin, 2014) to organise the embedded specification in order to extract the most pertinent information from this document. Several fields were included on the document summary form to assist the researcher in efficiently organising the information. These fields include the document's date, title, description, contents list, significance to the research field, and the researcher's reflections or considerations. Each field provides data that enhance the comprehension of the documents. Using this form, researchers could obtain an overview of the documents and expedite the content analysis process to determine whether teachers' assessment practises align with policy requirements (Yin, 2014). To maintain order and facilitate referencing, ascending-order labels were applied to the documents collected for this study. The first document, for instance, was labelled

“Document form 1” and designated “doc 1.” This labelling system improves the lucidity of document identification and facilitates the establishment of connections between different documents (Merriam, 1998).

Hence, to obtain access to the research site, the researchers sought authorization letters from two key entities: the Planning and Research Development Unit (PRD) of the Ministry of Education and the Penang State Education Department. This rigorous procedure was necessary to lay the foundation for our research endeavours. After an initial meeting with the school principals, a letter of information was sent to the selected research sites prior to the initial on-site visit. During this meeting, the researcher insured that the administrators were aware of the nature of this study and its adherence to strict confidentiality and data protection protocols. A document of informed assent was presented and signed with care, solidifying the commitment to protecting the participants’ data and privacy. The researcher then engaged with the ESL (English as a Second Language) instructors recommended by the school administrators. After being properly notified about the research project through the informed consent procedure and expressing a willingness to participate, these educators were contacted. Subsequently, a detailed schedule was established for conducting interviews and classroom observations, all of which were carefully coordinated to occur during regular school hours. This method was meticulously selected to capture the most authentic and representative educational experiences of ESL teachers and their students, thereby ensuring that our research was of the utmost significance.

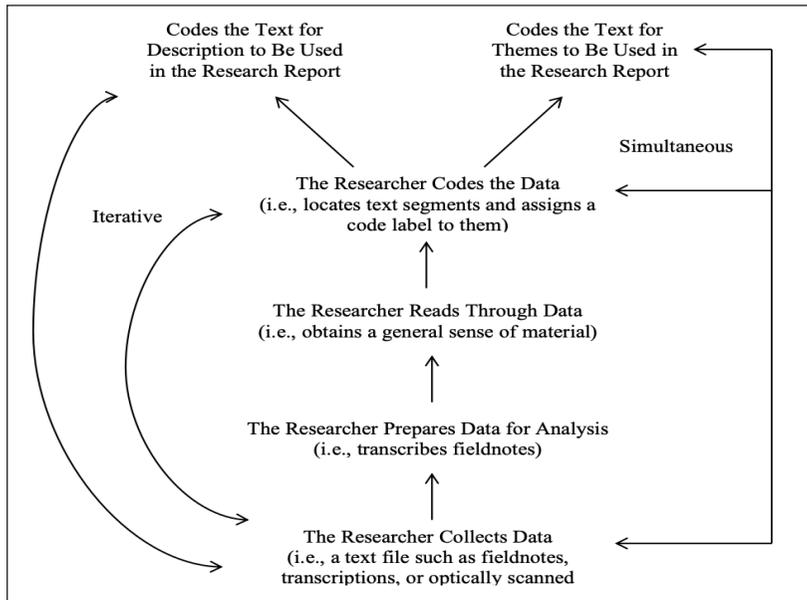
Data Analysis

To answer the research questions, patterns in teachers’ CBA practices were identified by analyzing observation and interview transcripts. Multiple readings and careful examinations of the transcripts led to a variety of codes regarding the CBA practices of teachers. Comparing the codes from each participant’s dataset yielded the major motifs of the study. As proposed by the Ministry of Education (2019) in the CBA Guidebook, these themes were grouped into three main categories (planning, implementing, and recording) to represent teachers’ assessment practices. In order to investigate the alignment of teachers’ assessment practices, a comparison to the curriculum standard document (SBELC) was also conducted, thus shedding additional

light on teachers' lived experiences in implementing the CBA within the context of Malaysian secondary school ESL classrooms. The findings of triangulated from the data collection methods (semi-structured interview, classroom observation, document analysis) were then analyzed through the thematic analysis procedures (Creswell, 2012, p. 237) as depicted in Figure 3.

Figure 3

The Qualitative Process of Data Analysis



The Atlas.ti programme for qualitative data analysis was utilised to manage the data. This form of QDA software assists researchers in storing and organising their data, enables them to assign labels or codes to their data, and makes it simpler to search the data for particular texts or words (Creswell, 2012). The purpose of these computer programs is not to generate an analysis, but rather to reduce the researcher's manual workload (Bryman, 2008). This indicates that the QDA programmes do not perform data analysis, but rather assist researchers in managing their data in a more organised manner. This procedure begins with the organisation and preparation of data for analysis. Upon data collection for this study, a separate file with separators was constructed and labelled for each participant.

Each participant's dossier included information from an interview, classroom observation, and curriculum excerpt. The subsequent phase in data analysis is to gain a broad understanding of the data (Creswell, 2012). Thus, the data was systematically investigated by participant. The data analysis started with each ESL teacher's data set, which included semi-structured interviews, classroom observations, and document analysis. The same process was repeated with each participant. The complexity of the research design necessitated the development of a system to ensure that the classification was done systematically and that no data were lost. Each code is preceded by "Planning", "Implementing", or "Reporting". Next, by tailoring the findings to the research questions together with the codes and sub-codes performed, the data was explored to formulate the themes in an endeavour to answer the research questions. (Creswell, 2012). Our approach to representing and reporting the findings was a narrative discussion in which the data analysis was discussed in depth. After the data was analysed, the interpretation of the findings was a crucial step in answering the research questions. The research findings were interpreted to form the overarching theme, which was then discussed in light of contemporary literature in order to derive the answers to the research questions.

The codes and themes identified in this study were subsequently subjected to Intercoder Reliability evaluation. The level of agreement over the coding of the procedure was assessed using Cohen's Kappa Threshold of Agreement, as presented in Table 2. This technique improves the reliability and validity of the findings presented in this study.

Table 2

Cohen's Kappa Threshold of Agreement

Values	Indication of Agreement
≤ 0	No agreement
0.01-0.20	Slight agreement
0.21-0.40	Fair agreement
0.41-0.60	Moderate agreement
0.61-0.80	Substantial agreement
0.81-1.00	Perfect agreement

The Trustworthiness of the Study

Despite of utilizing the triangulation method in reporting this study, an Intercoder Reliability (ICR) assessment was conducted on the discussed emerging codes and themes to ensure the validity and reliability of the findings. This evaluation sought to determine the coders' consistency and agreement with regard to the study's emerging codes and themes. By employing this rigorous evaluation technique, the study improves the reliability and validity of its findings, thereby bolstering its credibility. 15 codes were cross analysed by two independent raters, and an analysis of Cohen's Kappa was performed based on the analysis that arose in the coding frame from both raters. To create the first codes in the coding frame, Coder 1 segmented and labelled the data. The researcher repeated the same coding frame and handed the "clean" file to Coder 2 after the coding was finished and saved in the coded file. The transcription was then independently coded by Coder 2 using a framework that was comparable to the first. After the codes had been applied to the data unit, the degree of agreement between each code from both coders was then translated into a nominal form, with 0 denoting no agreement between raters and 1 denoting perfect agreement between raters. Since O'Connor and Joffe (2020) confirmed that Cohen's Kappa calculation runs between the value of 0 and 1, the rationale behind changing the codes into nominal form was to make the reliability calculation process easier. The SPSS analysis of the ICR assessment displays the results of Cohen's Kappa analysis. According to the results presented in the preceding table, the measure of coders' agreement was $k=0.65$, $p < 0.05$. The Cohen's Kappa analysis score was compared with the threshold indication of agreement (Landis & Koch, 1977) shown in Table 3.

Table 3

ICR Symmetric Measures

	Value	Asymptotic Standardized Error ^a	Approximate T ^b	Approximate Significance
Measure of Kappa Agreement	.65	.14	3.55	.00
N of Valid Cases	15			

a. Not assuming the null hypothesis.

b. Using the asymptotic standard error assuming the null hypothesis.

Based on the analysis depicted, this study revealed that there was “substantial agreement” between coders in the coding presented in this study as compared to the threshold of agreement (Table 2) in the preceding section. Hence, this accomplished the reliability of coding discussed in subsequent sections.

FINDINGS AND DISCUSSION

In accordance with the CBA Guidebook (2019) by the Ministry of Education, teachers’ CBA practices emerged in three primary dimensions: planning, implementing, and reporting the assessment. Consequently, this section explores the congruence of the teachers’ assessment practices with the policy’s aims and objectives while presenting these dimensions for elucidating the teachers’ assessment methods.

Planning the Assessment

According to the interview responses, before implementing classroom assessment, most teachers focused on defining their assessment and learning objectives. Teachers attempted to deconstruct the national curriculum standards into specific instructional objectives, which were categorized on two levels: the content and learning standards. In accordance with the national curriculum standards, the content standards were more generic in nature; however, the learning standards provide a more detailed description of the language skills, knowledge, and affective attitudes that students are expected to acquire. This is demonstrated by the actions of Teacher 2, who claimed to have consulted the curriculum document before administering an assessment to students.

“I like to decide on my assessment objectives first before assessing the students...erm... you know in this way it is easier for me to know what to focus during the evaluation process...” (T2)

In addition, Teacher 4 added that she utilised similar assessment planning strategies. She clarifies the instructional objectives in her practise by referring to the curriculum document, but she also determines these criteria based on the language proficiency of her students in the classroom. According to her, the curriculum standards

in the document were exceedingly advanced in comparison to the level of her students. Therefore, by adjusting the standards to meet the needs of her students' proficiency in the classroom, it has become simpler for her to assess what students "can do" in accordance with the CEFR descriptors.

... you know sometimes the document is just a reference for us, we can adjust whenever necessary, and this is what am I doing to make the assessment possible... (T4)

According to the claims made by Teachers 2 and 4, teachers' pay close attention to curriculum standard documents in order to ensure the standardisation of their assessment practises in accordance with curriculum objectives. Teachers were able to use the document to determine the focus of their classroom evaluation. A passage from the SBELC supported this assertion.

The SBELC document contains a mapping of the English Language Content and Learning Standards as well as CEFR-aligned pedagogical approaches.

(SBELC, 2017, p.18)

Although teachers were found to use the curriculum document (SBELC) as a guide when determining their assessment practices, there were instances in which teachers established lesson objectives by developing success criteria. Basically, the success criteria are derived from the learning objectives, but they are more specific. According to Rojon, McDowall, and Saunders (2015), the success criteria characterize the students' performance in terms of comprehension skills - what they will say, do, or produce to demonstrate that they have attained the learning objectives. Teachers made no allusion was made by the teachers to the curriculum document. Instead, teachers determine the success criteria based on the assigned task. This is evidenced by Teacher 1's practice, in which he claimed that the specification of learning objectives through success criteria made it simpler to evaluate the students.

to plan...erm...I mostly drawn it with the success criteria. Because for me the whatever decision I made through the success criteria is very useful. This way it provides me on how I can measure the students' performance during teaching and learning... (T1)

In another instance, teachers determined the criteria for success by referring to the materials and instruments from supplementary materials such as textbooks. For instance, Teacher 3 asserted that she consulted the textbook (Pulse 2), selected the assessment instrument, and then developed success criteria based on the task chosen.

...I created my own success criteria. This is driven from the textbook. This easiest way is select the task to be assessed to the students, then you can already decide the success criteria... (T3)

Based on the practices described in the planning stage of the CBA, it is evident that teachers were able to implement the assessment with discernment. By creating instructional objectives, teachers were deemed to have effective instructional practices for integrating assessment into the teaching and learning process in the classroom. In assessing the students in the classroom, teachers were able to ensure the achievement of developmental learning objectives. This is pertinent to the SBELC statement that the assessment should be based on the instructional objectives and be of a developmental nature instead of static. The following excerpt illustrates this matter.

An important consideration for teachers to bear in mind is the fact that the annual scheme of work should be developmental and not remain static. The scheme of work can be modified depending on the needs of the school and particularly the needs of the pupils based on their levels of proficiency and motivation.

(SBELC, 2017; p. 39)

This research confirms the findings of a study conducted in China by Yan et al. (2021). In planning the CBA, teachers in China utilised a similar strategy by breaking down the national curriculum standards into precise instructional objectives that were divided into three levels: term, unit, and lesson, leading to an effective implementation of the assessment in the classroom.

Implementing the Assessment

This section describes the teachers' practises for implementing the CBA during classroom teaching and learning. The data presented

in this section will be derived from interviews with and classroom observations of the teachers. This section will also examine the compatibility of their assessment procedures with the policy's requirements. On the basis of the collected data, it can be concluded that teachers employed various assessment strategies and methods depending on the skills to be assessed (productive or receptive skills) in the classroom. Thus, three themes (written, observation and oral) emerged representing teachers' assessment strategies in enacting the CBA.

Written Assessment

Classroom assessments in written form involve students showcasing their comprehension of a specific topic through writing (Tayyebi et al., 2022). This assessment method is commonly employed by teachers to gauge students' progress, knowledge retention, and their ability to effectively articulate their thoughts, as cited in the same source (Tayyebi et al., *ibid*). Written assessments within the classroom can take various forms, including essays, short-answer questions, multiple-choice queries, and problem-solving tasks. In this study's context, the assessment process aligned with textbook chapters and made use of task sheets. For example, Teacher 3 and Teacher 1 shared her practice of using the textbook for student assessments, conducting "chapter-based assessments" on a monthly basis. She believed that this strategy would allow her to gather ongoing assessment data to track her students' knowledge growth.

For me, textbook has always the best assessment tool. I always used this material and assess the students based on the chapters in the textbook...erm...usually once in a month maybe? (T3)

...it's easier to use the materials and instrument from the textbook...you know...it's more guided in a way... (T2)

In a different practice, Teacher 4 also used the task sheet to administer the written assessment to students. In this context, teachers use supplementary materials, such as the exercise or activity book, to administer assessments to students. For instance, Teacher 4 and Teacher 3 claimed that the written assessment was also administered this way because it was simpler for them to determine how students

responded to the questions. The questions on the task sheet are based on the examination questions of the students.

I used the worksheet and design the questions accordingly to their exam format. For me it is easier to train students to answer exam like questions...you know...to get them used to the questions... (T4)

...for me drilling students with the exam situation questions can help them to improve... (T3)

Although this study revealed that the textbook and task sheets were heavily utilised in the implementation of CBA, it is important to note that this practice did not deviate from the traditional exam-centric approach. Upon close inspection of Teacher 3's classroom assessment on reading skills for the intermediate learners, it was evident that essential scaffolding was absent, and she placed a significant emphasis on the students' ability to independently complete assessment tasks. As a result, a large number of students hurriedly attempted to complete their assignments within the allotted time frame, resulting in confusion regarding the lesson objectives they were to achieve. Furthermore, after completing the assessment procedures, Teacher 3 promptly collected her students' grades. This approach stands in stark contrast to the recommendations put forth in the SBELC excerpt. The guidelines from SBELC advocate for a less stressful assessment environment and emphasize that assessments should not be primarily used for comparing students' abilities in earning high grades. The SBELC approach promotes a more holistic and less grade-centric perspective on assessment, which diverges from the hurried and grade-focused practices observed in Teacher 3's classroom.

In order for interaction between teachers and students, students and students, and students and learning materials to be meaningful, teachers must establish a fun, pressure-free, and modest classroom environment. (CBA Guidebook, 2019; p.39)

In addition, while the assessment focused predominantly on score accumulation, an additional classroom observation of Teacher 4's instructional practises revealed a lack of appropriate differentiation strategies. During this classroom observation on the writing skills,

students were subjected to the same assessment tools and queries derived from the textbook and assignment sheet, irrespective of the diverse abilities within the student population. This approach is consistent with the prevalent culture that places a premium on examinations and may put pressure on students categorised as ‘low performers’ to perform at their highest level. This practise is in stark contrast to the SBELC’s recommendations, which emphasise the significance of educators employing differentiation strategies in their assessment procedures. Such strategies promote a more inclusive and equitable evaluation process by ensuring that students are evaluated based on their individual “can do” abilities.

It is extremely important that teachers are aware that a pupil’s proficiency in English is not fixed and because a pupil is currently less proficient in English, this does not mean in any way that this will remain true throughout a pupil’s secondary education. It is important that all pupils are given equal opportunity to develop over time and teachers should ensure that all their pupils are challenged and given equal opportunity for development.

(SBELC, 2017; p. 20)

Although the CBA was intended to move away from this culture, the fact that a great deal of emphasis was placed on the practice of using the task sheet, particularly the accumulation of marks during the exam, suggests that teachers’ practices were still influenced by the exam-oriented culture. Similar results were discovered in a study conducted by Yigzaw (2013) in Ethiopia, where formal testing was still a prevalent practice among teachers implementing classroom-based assessment. It was claimed that it was unethical to use feedback such as grades and general commendation to determine whether students had mastered the language knowledge objective.

Observation Based Assessment

Moreover, teachers also assessed students’ language skills, particularly their listening and speaking abilities, through classroom observations. As indicated by Puspita and Suyatno (2020), this strategy capitalises on the teachers’ ability to collect real-time data regarding their students’ performance, conduct, and attitudes. This observational technique was utilised by teachers to obtain direct insight into student

learning and classroom interactions. Within the scope of this study, teachers implemented collaborative student strategies, most notably peer evaluation techniques. This was exemplified in Teacher 2's statement that her students were instructed to work in teams. She acknowledged that integrating multiple language skills within a single lesson activity makes peer evaluation within group activities simple and effective. Through this method, students not only improved their language skills but also had the opportunity to assess the performance of their classmates.

I favour group projects because they incorporated a collaborative learning strategy. So, whenever they engage in group work, they listen to the presentation of their peers... (T2)

Moreover, in a similar practice, Teacher 1 also claimed that she made use of the technique of question and answer during the peer evaluation. She engaged the students through the collaborative work by encouraging her students to post questions to their peer. She added that through this method she has been able to engage the students' centred approach during the teaching and learning process in the classroom.

During the peer evaluation, I also ask the students to ask the questions to their friends. Through this way I can create the nature of students centredness in my pedagogy... (T1)

However, the observational method used to implement CBA did not meet the policy's requirements. For example, during an observation of Teacher 2's classroom, the process of peer evaluation was not well executed because the instructional objectives were not explicitly communicated to the students in the classroom. The students did not understand what constituted high-quality work. As a result, they tended to evaluate their own and others' performance summatively. When one student was asked to evaluate his own handwriting, he could only provide a general assessment of his performance by stating, "I have not done a good job." Although classmates were asked to provide feedback on the strengths and weaknesses of a student's oral presentation, they tended to overlook the presentation's strengths and focus on its weaknesses. Consequently, the constructive feedback provided during peer evaluation was poorly delivered. A comparable

scenario was also observed in Teacher 1's lower secondary speaking class. The method of utilizing a question-and-answer technique did not effectively involve all students in active participation during the class. During the observation, it became apparent that many students were not actively engaged in the learning process, as the teacher primarily directed the instruction toward the advanced learners, inadvertently neglecting the intermediate and beginner groups. This circumstance hindered the assessment of students' performance due to this instructional approach. This is in contrast to the information in the document, which stated that the assessment should inform the students' developmental feedback. In this sense, CBA Guidebook also supported that the only way to effectively administer classroom evaluation is to inform students of the success criteria and instructional objectives prior to executing the assessment.

Teachers must understand the educational objectives outlined in curriculum documents. Each subject's curriculum incorporates knowledge, skills, and values. To evaluate the level of pupil mastery, teachers must comprehend the purpose of the secondary school curriculum framework as well as the Content Standards, Learning Standards, and Performance Standards for their specific subjects.

(CBA Guidebook, 2019; p. 39)

The study by Yan et al. (2021) likely highlights the significance of explicit learning objectives and the impact their absence has on peer assessment. Due to the teachers' inability to clarify explicit learning objectives to students throughout the instructional session, peer assessment during classroom instruction failed. This issue has impeded the students' ability to comprehend the requirements they must meet to complete the lesson's objective and the evaluation being conducted. To address this issue, it is crucial that teachers articulate learning objectives at the start of each lesson. This can be accomplished via verbal explanations, written instructions, or visual aids that provide students with a road map.

Oral Assessment

In addition to a written evaluation, ESL teachers also used an oral evaluation to implement the CBA. This study found that oral

presentations have become the most common method for teachers to administer oral assessments systematically. The presentation served as a formal mechanism for instructors to assess students' classroom performance. Students were instructed to compose their responses according to the mission requirements. For example, Teacher 1, who formally evaluated his students using oral presentations, stated that this method was beneficial for capturing the students' performance during the actual classroom session. To monitor the performance of his pupils in the classroom, she incorporated oral presentations and peer evaluations into her evaluation methods. Throughout the teaching and learning, students were asked to present a topic, and during the evaluation, they were required to assess the performance of their classmates.

...usually, I will ask them to present because it is easier for me to evaluate for their performance and sometimes ask the other students to give feedback about their friends' performance... (T1)

In addition, Teacher 3 who also utilized the oral assessment method reported that the evaluations were conducted alongside the criteria she created himself. In this situation, she was equipped with a rubric as a reference for her to evaluate the students' performance during the assessment. One of the primary purposes of the checklist is to facilitate the process of evaluating students, as it clarifies the aspects on which he should concentrate when evaluating students.

Oh ya...I also use the checklist of my own to assess the students' performance in this evaluation method. For me it is easier as the checklist and rubric guides me on the aspect that I need to look for in assessing the students. In short, it's structured in a way... (T3)

While the teacher professed to have extensively employed rubrics and checklists during the oral assessments, it was evident that the consistent application of rubrics in the assessment process was lacking, given the diverse array of students in their classroom. Using identical rubrics and checklists for the oral assessments has had adverse effects on the students' potential, particularly those categorized as 'low performers,' making it challenging for them to showcase their true abilities. This was clearly observed during the classroom assessment of Teacher

1, where the assessment process was disrupted due to the rubric's disproportionate focus on elements that these struggling students find particularly challenging, thus impeding their ability to demonstrate their full potential. A parallel situation was also witnessed during the classroom observation of Teacher 3. This research uncovered that employing a rubric that does not align with the students' varying capabilities has compromised the integrity of the classroom assessment procedures. Teacher 3, who imposed exceedingly high criteria for her beginner students, disrupted the flow and direction, making it challenging for these students to convey their language learning abilities effectively. In summary, the misalignment between rubric use and the diverse needs of students in these classrooms has raised concerns about the fairness and effectiveness of the assessment process. In fact, according to Simpson and McKay (2013), using a rubric that is not appropriately tailored to the individual requirements and abilities of students can result in an inaccurate assessment and inhibit their development. Stanford and Reeves (2009) research elucidated the similar circumstance regarding this issue. The study found that designing a rubric that is not authentic to the classroom environment jeopardised the assessment of students in the United States because they were unable to demonstrate knowledge that adheres to the rubric's requirements.

Reporting the Assessment

This section discusses the practises of teachers following the implementation of CEFR-aligned CBA. This section describes how teachers record and analyse the mastery level of their students after implementing classroom assessments by focusing on how the assessment was recorded and documented for the benefit of other stakeholders. Consequently, two major themes (the transit form and standard setting) emerged to characterise the teachers' practices in classroom assessment after its implementation.

The Transit Form

This study revealed that the recording template referred to as the transit form was utilised substantially in the post-implementation of classroom evaluation. The transit form is a record of the student's progression over time. It typically includes a list of learning objectives or standards and spaces for teachers to record the results or grades

for each assessment. Throughout the school year, teachers must use this form to monitor student performance and identify areas in which students may need additional support during the intervention programme. This is supported by the assertion of Teacher 2, who used the transit form to monitor and record her students' competency levels over the course of multiple years of assessment administration. She was given a transit form with two primary columns representing data collected during the first and second halves of the year. As a result of implementing and documenting the CBA twice a year, teachers at her school were able to compare and track student progress during classroom assessments.

In that file, we must include the transit form, a list of student names, and for years we will do twice PBD (CBA), so we must have a list of student names and two columns for the first half of the year and the second half of the year. I believe they are more interested in the progress of the students throughout the year than in the final result. (T2)

In addition, Teacher 3 was also provided with a recording template to document her students' performance on the course assessment. In her school, the transit form was regarded as the official document for administrators' reference. Also, it entailed accumulating teachers' personal records regarding students' progress, according to her account. The transit form used at her school was divided according to the skills based on the content and learning standards of the curriculum documents. This alignment can guarantee that teachers are assessing the essential skills and knowledge students must acquire. Therefore, they were required to implement a continuous classroom assessment process by evaluating students' performance against the SBELC content and acquiring standards-aligned skills.

We do possess the sheet or template supplied to me by my panel head. It's a template provided by the school. The template is subdivided according to the talents I intend to evaluate. It is for my own records that I periodically refer to their progress. (T3)

It appears, based on the practices described previously, that the implementation of CBA has had unintended effects on how

teachers approach classroom assessment. Teachers may place a disproportionate emphasis on recording and documenting data for administrative purposes instead of developing effective assessment practices to support student learning. The transit form and template were considered crucial documents teachers should prepare in class. This is evident from the claims made by Teacher 3, who expressed dissatisfaction with administering the CBA because she perceived this assessment to require a laborious procedure, particularly in compiling the data.

For me, CBA is a complicated process. It demands a lot of tedious work especially on the part when we have to key in and preparing all the data. This is all the demands from the higher up... (T3)

In addition, she asserted that the emphasis on preparing assessment data had diverted her attention away from developing an efficient classroom assessment procedure. This pertains to the type of information the teacher would provide to support her judgment that the evaluation decision is accurate, making it difficult to administer classroom-based assessments.

They have always commented on our data, some part is not reliable... some part is okay...erm... and at some point, I have to change the data in order to suit the needs and demands of them... (T3)

This is consistent with the claim made by Smith (2003), who stated that the implementation of classroom assessment has been found to be highly focused on developing grading data, which at times diverts the primary focus from exploring students' learning capabilities. This has created a contentious site, as the implementation of CBA should not be heavily concentrated on assessment data, but rather on how teachers were able to be informed about the growth of students throughout the learning process.

Standard Setting

This study also revealed that, after implementing CBA, teachers conducted standard-setting procedures. According to Wills et al. (2016), standard setting is a practice where teachers share and develop

their understanding of what learning looks like by comparing various types and levels of student work to formal standards and success criteria. According to Kampa et al. (2019), the term standard-setting encompasses a variety of consensual approaches that involve teachers setting discrete cut scores and comparing them to proficiency scales. In brief, it is the practice of identifying the aspects or criteria of classroom assessment that can assist teachers in exercising their professional judgement. This process involves making consistent decisions when assessing students' mastery level, which can narrow the professional judgement disparity between teachers. Wiggins and McTighe (2005) asserted that this stage is essential for ensuring the evaluation's validity and reliability. In the context of this study, standard setting was conducted to ensure that teachers are awarding students consistent grades in their classrooms. This is supported by Teacher 1's assertion that the process was conducted during the post-implementation phase of CBA as part of the assessment's quality control to ensure that students' mastery levels are uniform across all teachers.

What we did in the panel is to do the standard setting. We present our assessment data and compare to the others. You know, for the sake of having an even distribution of data across teachers in the panel. So that, no issue this teacher is strict and that teacher is lenient. (T1)

In a different practise, Teacher 4 who utilised the standard-setting procedures explained that during the procedures, teachers in the panel meet in a Professional Learning Community (PLC) group to determine the cut scores of the grade students earned on classroom assessments and subdivide it into the determined performance levels.

...the teachers will meet to set the standard for each ensemble in accordance with the grade obtained by the students. So, following the discussion, I recorded my students' mastery level...aaa... based on the standard that all teachers have agreed to establish... (T4)

This approach concluded that establishing standards involves subjective evaluations, making it a highly contentious endeavour. It was discovered that teachers did not utilize the SBELC performance standards to guide their professional discernment when carrying out their standard-setting procedures, indicating that

they were disconnected from the policy's requirements. This starkly contrasts claims made in the CBA Guidebook (2019), which mandates that teachers to consult the curricular standard document when making professional decisions, as evidenced by the following excerpt.

The decision on the students' mastery level is depends on the teacher's autonomy and it should be accomplished by considering the subject's Curriculum and Assessment Standard Document (DSKP) requirements and specifications.

(CBA Guidebook, 2019; p.39)

Teachers' heavy reliance on the accumulation of student marks and scores in the implementation of CBA indicates that their practises are diverted from the policymakers' primary goals for classroom assessment. The excessive reliance of teachers on accumulating students' test scores demonstrates that their practises are identical to the previous assessment system, which was exam centric. In this sense, the implementation of CBA did not run away from the exam-centric culture by not comparing student performance based on their assessment scores. The teachers' practices described earlier contrast starkly to the specification embedded in the CBA Guidebook (2019) which states that,

All assessment information is not intended to be used for student comparison or competition. Instead, the information should be utilised to assist the school and parents in planning follow-up actions aimed at enhancing student learning mastery and achievement.

(CBA Guidebook, 2019; p.5)

In addition, Brown and Harris (2016) confirmed that an excessive reliance on the accumulation of marks during formative assessment can detract from the productiveness of the feedback. They added that in order for feedback to be productive and developmental in nature, the collection of numerical data cannot indicate how well students have mastered the lesson; rather, it is necessary to retrieve 'information' that teachers can use to enhance their teaching and learning instruction. A grade or number conveys very little information about the feedback, particularly regarding how to improve and address problem areas (Brown & Harris, 2016).

CONCLUSION

In conclusion, this study revealed that Malaysian ESL teachers shown competence in instructional and pedagogical techniques, although their evaluation approaches did not conform to the norms of CBA. The pervasive examination-driven culture has pushed teachers to prioritize grades over language proficiency, impeding the process of learning. The teachers' deficiency in assessment literacy resulted in ineffectual evaluations and students' perplexity over objectives. In this sense, the implementation of the CBA did not meet its objectives, highlighting a discrepancy between the curriculum and teachers' assessment methods. To tackle this issue, it is recommended to review the CBA policy, offer teachers evaluation literacy training, and prioritize experiential learning. Professional development should integrate role-playing, case studies, and group discussions customized to address the individual requirements of teachers. The objective of this strategy is to improve comprehension and implementation of the recently implemented assessment policy in their teaching methodologies. On top of that, teachers must also acknowledge the significance of classroom assessment in educational reform, as it allows for the monitoring of progress and the formulation of informed instructional decisions. The implementation of continuous formative assessments, as highlighted in the CEFR-aligned CBA, provides a more thorough evaluation compared to the prior system that focused mostly on exams. Teachers must understand these distinctions in order to apply successful instructional strategies that are in line with the new policy, offering accurate feedback and enhancing student achievements. Thus, to achieve successful implementation, the curriculum document must clearly articulate the concept of classroom assessment, provide specific and concrete examples, and provide guidance to teachers on how to integrate it into their lesson planning. Providing explicit guidance on the many forms of assessments, their timing, interpretation, and utilization of data will guarantee uniformity and facilitate efficient training, ultimately improving student outcomes.

ACKNOWLEDGMENT

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

REFERENCES

- Andrade, H. L., & Cizek, G. J. (2010). *Handbook of formative assessment*. Routledge.
- Arumugham, K. S. (2020). Kurikulum, pengajaran dan pentaksiran dari perspektif pelaksanaan pentaksiran bilik darjah. *Asian People Journal*, 3(1), 152–161. <https://doi.org/10.37231/apj.2020.3.1.175>
- Azli, N., & Akmar, A. (2019). Implementation of CEFR-Aligned assessment tools in Malaysian ESL classroom. *Asia Proceedings of Social Sciences*, 4(2), 7–10. <https://doi.org/10.31580/apss.v4i2.688>
- Baker, L. (2006). Observation: A complex research method. *Library Trends*, 55(1), 171–189. <https://doi.org/10.1353/lib.2006.0045>
- Black, P. (2000). Research and the development of educational assessment. *Oxford Review of Education*, 26(3-4), 407–419. <https://doi.org/10.1080/713688540>
- Black, P. (2015). Formative assessment – an optimistic but incomplete vision. *Assessment in Education: Principles, Policy & Practice*, 22(1), 161–177. <https://doi.org/10.1080/0969594x.2014.999643>
- Black, P., & Wiliam, D. (2003). “In praise of educational research”: Formative assessment. *British Educational Research Journal*, 29(5), 623–637. <https://doi.org/10.1080/0141192032000133721>
- Brown, G. T., & Harris, L. R. (2016). *Handbook of human and social conditions in assessment*. Routledge.
- Bryman, A. (2008). *Social research methods* (Third ed.). Oxford: Oxford University Press.
- Chen, Q., May, L., Klenowski, V., & Kettle, M. (2014). The enactment of formative assessment in English language classrooms in two Chinese universities: Teacher and student responses. *Assessment in Education: Principles, Policy & Practice*, 21(3), 271–285.
- Cheng, L., Rogers, W. T., & Wang, X. (2008). Assessment purposes and procedures in ESL/EFL classrooms. *Assessment & Evaluation in Higher Education*, 33(1), 9-32.
- Chin, H., Thien, L. M., & Chiew, C. M. (2019). The reforms of national assessments in Malaysian education system. *Journal of Nusantara Studies*, 4(1), 93-111. <https://doi.org/10.24200/jonus.vol4iss1pp93-111>

- Creswell, J. W. (2002). *Educational research: Planning, conducting, and evaluating quantitative and qualitative research*. Pearson Education International.
- Creswell, J. W. (2012). *Educational research: Planning, conducting and evaluating quantitative and qualitative research* (Fourth ed.). Boston: Pearson Education.
- Curriculum Development Division. (2017). *Dokumen Standard Kurikulum dan Pentaksiran. Pendidikan Moral Tingkatan Dua*. Putrajaya: Bahagian Pembangunan Kurikulum, Kementerian Pendidikan Malaysia.
- Ghavifekr, S., Kunjappan, T., Ramasamy, L., & Anthony, A. (2016). Teaching and Learning with ICT Tools: Issues and Challenges from Teachers' Perceptions. *Malaysian Online Journal of Educational Technology*, 4(2), 38–57. <https://eric.ed.gov/?id=EJ1096028>
- Hajmi, H. H., & Jamaludin, K. (2022). Kesediaan guru dalam pelaksanaan pentaksiran bilik darjah di sekolah rendah. *Jurnal Dunia Pendidikan*, 4(4), 119–125. <https://myjms.mohe.gov.my/index.php/jdpd/article/view/20751>
- Harlen, W. (2007). *The quality of learning*. Routledge.
- Hopfenbeck, T. (2018). Classroom assessment, pedagogy and learning – twenty years after Black and Wiliam 1998. *Assessment in Education: Principles, Policy & Practice*, 25(6), 545–550. <https://doi.org/10.1080/0969594x.2018.1553695>
- Kampa, N., Wagner, H., & Köller, O. (2019). The standard setting process: Validating interpretations of stakeholders. *Large-scale Assessments in Education*, 7, 1-25. <https://doi.org/10.1186/s40536-019-0071-8>
- Landis, J. R., & Koch, G. G. (1977). An application of hierarchical kappa-type statistics in the assessment of majority agreement among multiple observers. *Biometrics*, 363-374.
- Lembaga Peperiksaan. (2012). *Panduan Pengurusan Pentaksiran Berasaskan Sekolah*. Kementerian Pendidikan Malaysia.
- Leong, W. S. (2014). Knowing the intentions, meaning and context of classroom assessment: A case study of Singaporean teacher's conception and practice. *Studies in Educational Evaluation*, 43, 70–78. <https://doi.org/10.1016/j.stueduc.2013.12.005>
- Liu, J., & Xu, Y. (2017). Assessment for learning in English language classrooms in China: Contexts, problems, and solutions. *Innovation in language learning and teaching: The case of China*, 17-37.

- Lumadi, M. W. (2013). Challenges besetting teachers in classroom assessment: An exploratory perspective. *Journal of Social Sciences*, 34(3), 211-221.
- Mat Yusoff, S., Razak, R. A., Chin, H. L., & Marzaini, A. F. (2023). Exploring teachers' conceptions of assessment: A quantitative study in a secondary school setting. *Journal of Research, Policy & Practice of Teachers and Teacher Education*, 13(2), 59-72.
- Marnizam, F., & Ali, S. (2021). Evaluation of the implementation of classroom assessment (PBD) among primary school mathematics teachers. *Jurnal Pendidikan Sains Dan Matematik Malaysia*, 11(2), 81-94. <https://doi.org/10.37134/jpsmm.vol11.2.7.2021>
- Merriam, S. B. (1998). *Qualitative research and case study applications in education*. Jossey-Bass Publishers.
- Merriam, S. B., & Tisdell, E. J. (2016). *Qualitative research: A guide to design and implementation* (4th ed.). Jossey-Bass, Cop.
- Ministry of Education (2017). *Standard Based English Language Curriculum for Secondary School*. Kementerian Pendidikan Malaysia.
- Ministry of Education (2019). *Classroom Based Assessment Guidebook*. Kementarian Pendidikan Malaysia.
- Mohamad Hanapi, M. H., Zakaria, N., & Muner, S. (2021). Readiness level of primary school teachers in Klang district, Selangor in the implementation of in-class assessment from the aspect of knowledge. *International Journal of Modern Education*, 3(9), 01–08. <https://doi.org/10.35631/ijmoe.39001>
- Mohamad Marzaini, A. F., Haslee Sharil, W. N. E., Supramaniam, K., & Mat Yusoff, S. (2023). Evaluating teachers' assessment literacy in enacting cefr-aligned classroom-based assessment in Malaysian secondary schools ESL classroom. *International Journal of Academic Research in Progressive Education and Development*, 12(1), 661-675. <https://doi.org/10.6007/ijarped/v12-i1/15691>
- Mohd Isa, A., Mydin, A.-A., & Abdullah, A. G. K. (2021). Road to school transformation 2025: A systematic literature review on teacher autonomy in Malaysia. *International Journal of Academic Research in Progressive Education and Development*, 10(3), 34-45. <https://doi.org/10.6007/ijarped/v10-i3/10710>
- O'Connor, C., & Joffe, H. (2020). Intercoder reliability in qualitative research: Debates and practical guidelines. *International Journal of Qualitative Methods*, 19, 228–232. <https://doi.org/10.1177/1609406919899220>

- Panadero, E., Fraile, J., Fernández Ruiz, J., Castilla-Estévez, D., & Ruiz, M. A. (2019). Spanish university assessment practices: examination tradition with diversity by faculty. *Assessment & Evaluation in Higher Education*, 44(3), 379-397.
- Pastore, S. (2023). Italian teachers' conceptions of large-scale assessment: What are the implications for assessment literacy?. *Teacher Development*, 27(2), 153-171.
- Patallion Jr, A. P. (2016). The implications of learning theories to assessment and instructional scaffolding techniques. *American Journal of Educational Research*, 4(9), 695-700. <https://doi.org/10.12691/education-4-9-9>
- Patton, M. Q. (2015). *Qualitative Research & Evaluation Methods: Integrating Theory and Practice* (4th ed.). Sage Publications.
- Pellegrino, J. W., DiBello, L. V., & Goldman, S. R. (2016). A framework for conceptualizing and evaluating the validity of instructionally relevant assessments. *Educational Psychologist*, 51(1), 59-81. <https://doi.org/10.1080/00461520.2016.1145550>
- Puspita, R. R., & Suyatno, S. (2020). Teacher observation assessment for primary education. *ETERNAL (English Teaching Journal)*, 11(2), 78-82. <https://doi.org/10.26877/eternal.v11i2.7565>
- Rahman, N. H. (2014). *From curriculum reform to classroom practice: An evaluation of the English primary curriculum in Malaysia* [PhD Thesis].
- Rethinasamy, S., Ramanair, J., & Chuah, K.-M. (2021). English Medium Instruction at Crossroads: Students' Voice and Way Forward. *International Journal of Academic Research in Business and Social Sciences*, 11(14), 109-123. <https://doi.org/10.6007/ijarbs/v11-i14/8533>
- Rojon, C., McDowall, A., & Saunders, M. N. (2015). The relationships between traditional selection assessments and workplace performance criteria specificity: A comparative meta-analysis. *Human Performance*, 28(1), 1-25. <https://doi.org/10.1080/08959285.2014.974757>
- Sidhu, G. K., Kaur, S., & Chi, L. J. (2018). CEFR-aligned school-based assessment in the Malaysian primary ESL classroom. *Indonesian Journal of Applied Linguistics*, 8(2), 452-463. <https://doi.org/10.17509/ijal.v8i2.13311>
- Simpson, Z., & McKay, T. M. (2013). Assessment rubrics: Artefacts that speak in tongues? *Per Linguam*, 29(1), 15-32. <https://doi.org/10.5785/29-1-539>
- Smith, J. K. (2003). Reconsidering reliability in classroom assessment and grading. *Educational Measurement: Issues and Practice*, 22(4), 26-33. <https://doi.org/10.1111/j.1745-3992.2003.tb00141.x>

- Stanford, B., & Reeves, S. (2009). Making it happen: Using differentiated instruction, retrofit framework, and universal design for learning. *Teaching Exceptional Children Plus*, 5(6), n6.
- Suppian, Z., Mohd Isa, N. J., & Govindasamy, P. (2020). Penilaian sendiri guru pelatih terhadap tahap kemahiran pentaksiran bilik darjah (PBD). *Jurnal Dunia Pendidikan*, 2(4), 98–106. <https://myjms.mohe.gov.my/index.php/jdpd/article/view/11774>
- Taneri, P. O. (2016). Exploring teacher candidates' views on educational reform. *Athens Journal of Education*, 3(1), 19–31. <https://eric.ed.gov/?id=EJ1216499>
- Tayyebi, M., Abbasabady, M. M., & Abbassian, G.-R. (2022). Examining classroom writing assessment literacy: A focus on in-service EFL teachers in Iran. *Language Testing in Asia*, 12(1), 1-20. <https://doi.org/10.1186/s40468-022-00161-w>
- Tee, D. D., & Ahmed, P. K. (2014). 360 degree feedback: An integrative framework for learning and assessment. *Teaching in Higher Education*, 19(6), 579–591. <https://doi.org/10.1080/13562517.2014.901961>
- Wiggins, G., & McTighe, J. (2005). *Understanding by design* (2nd ed.). Association for Supervision and Curriculum Development.
- Willis, J., Adie, L., & Klenowski, V. (2013). Conceptualising teachers' assessment literacies in an era of curriculum and assessment reform. *The Australian Educational Researcher*, 40(2), 241–256. <https://doi.org/10.1007/s13384-013-0089-9>
- Yan, Q., Zhang, L. J., & Cheng, X. (2021). Implementing classroom-based assessment for young EFL learners in the Chinese context: A case study. *The Asia-Pacific Education Researcher*, 30(6), 541–552. <https://doi.org/10.1007/s40299-021-00602-9>
- Yeh, L. H. (2021). Challenges faced by visual art education with the existence of classroom based assessment in primary school within district of Hilir Perak. *Kupas Seni*, 9(1), 46–57. <https://doi.org/10.37134/kupasseni.vol9.1.5.2021>
- Yigzaw, A. (2013). High school English teachers' and Students' perceptions, attitudes and actual practices of continuous assessment. *Educational Research Review*, 8(16), 1489-1498.
- Yin, R. K. (2014). *Case study research: Design and methods* (5th ed.). Sage Publications.
- Zhao, X., Heuvel-Panhuizen, M. van den & Veldhuis, M. (2017, February 1). *Using classroom assessment techniques in Chinese primary schools: Effects on student achievement*. Hal. science. <https://hal.science/hal-01949253/>